ChinaXiv合作期刊
图书情报工作
第64卷 第2期 2020年1月

免费与付费在线问答社区用户参与行为的比较研究*

■ 齐云飞[1] 赵宇翔[2] 刘周颖[3] 孙晓宁[4] 朱庆华[3]
[1] 郑州航空工业管理学院信息管理学院 郑州 450015 [2]南京理工大学经济管理学院 南京 210094
[3] 南京大学信息管理学院 南京 210023 [4]山西财经大学信息学院 太原 030006

摘 要：[目的/意义]作为目前主流的在线问答社区模式，免费和付费在线问答社区为用户提供了不同的信息交互场景，能够满足差异化的用户需求。本文通过对两类社区用户参与行为的比较，深入揭示了他们之间的行为差异，对两类社区服务模式的优化提供了对策和建议。[方法/过程]采用列联表分析、非参数检验和负二项回归等方法对免费和付费在线问答社区用户的参与特征和交互影响关系进行了分析。[结果/结论]研究发现免费和付费在线问答社区用户在信息需求、参与特征和交互影响关系上存在诸多差异，其中免费在线问答社区具有最多的观点型问题，付费在线问答社区具有最多的建议型问题；免费提问者倾向于用抽象、概括的方式提出问题，付费提问者则倾向于提供详细的背景信息；问题长度负向影响免费回答的数量和长度，正向影响付费回答的长度；问题长度、回答长度正向影响免费和付费评价数量。本文丰富了对在线问答社区用户参与行为的研究，为免费和付费两类社区服务模式优化提供了一定的管理建议。

关键词：在线问答社区 付费问答平台 知识付费 用户信息交互
分类号：G250
DOI：10.13266/j.issn.0252－3116.2020.02.012

1 引言

在线问答社区是用户在自身需求和收益驱动下，以问答的方式进行信息生产、交换和实践的平台[1]，其具有鲜明的社会化特征和多种类型的群体协作模式，是汇聚集体智慧、优化知识管理和创新的重要工具[2]。在长期发展过程中，互联网领域涌现出许多知名的在线问答社区，2002年，谷歌公司推出了第一个网络知识付费社区——Google Answers。之后，Yahoo！Answers、百度知道、知乎等国内外知名的在线问答社区开始兴起，他们通过提供免费的信息交互渠道吸引了大量用户的使用。2016年，在共享经济、互联网+等新理念和新技术的推动下，我国掀起了一场知识付费的浪潮，以值乎、分答、丁香医生为代表的新一代付费在线问答社区如雨后春笋般涌现。2018年，上述平台的发展开始遭遇瓶颈，曾经火热的知识付费市场逐渐降温，业界和学界也开始对这一新兴模式进行反思。

作为目前两种主流的在线问答社区模式，免费和付费在线问答社区在平台模式、参与方式和激励机制等方面存在着较大差异。以Yahoo！Answers、知乎、百度知道为代表的免费在线问答社区采取的是社会化问答模式，用户基于个人兴趣进行信息搜索、分享和评价，社区主要通过荣誉、积分等虚拟奖励激励用户参与[3-4]。以值乎、分答等为代表的新一代付费在线问答社区采取的是专家问答模式，用户通过支付费用向社区专家提问。与Google Answers不同的是，上述社区均采取了新的交互形式和收益分配规则，无论是提问者还是回答者都有机会获得金钱奖励[5-6]。免费和付费在线问答社区功能、模式上的差异必然会对用户的参与行为产生影响[7]，他们会根据个人需求选择访问免费或是付费在线问答社区[8-10]，并推动社区形成独特的信息交互方式[11-13]。本研究希望通过对用户参

* 本文系国家自然科学基金项目“基于科研众包模式的公众科学项目运作与管理机制研究”（项目编号：71774083）和教育部人文社会科学研究青年基金项目“任务复杂性情境下的用户社交与协同信息搜寻行为机制研究”（编号：18YJC870018）研究成果之一。

作者简介：齐云飞（ORCID：0000－0002－8758－2837），讲师，博士，E－mail：yfqi2015@qq.com；赵宇翔（ORCID：0000－0001－9281－3030），教授，博士生导师；刘周颖（ORCID：0000－0001－5201－500X），博士研究生；孙晓宁（ORCID：0000－0003－1759－2543），讲师，博士；朱庆华（ORCID：0000－0002－4879－399X），副院长，教授，博士生导师。

收稿日期：2019－04－16 修回日期：2019－07－31 本文起止页码：105－115 本文责任编辑：王传清

与行为数据的分析,明确他们在免费和付费在线问答社区环境下信息需求类型、参与行为特征和交互影响关系上的差异,探索凸显免费和付费社区服务特色的优化方法,为两类社区的差异化发展提供对策、建议。

目前,针对在线问答社区的研究主要集中在对免费用户参与动因[14-17]或付费用户购买策略的分析[18-23],很少有研究对这两类社区用户的参与行为进行比较,对免费和付费在线问答模式的特点缺少系统、深入地了解。因此,本研究首先对与免费和付费在线问答社区、在线问答社区用户参与行为相关的研究进行梳理;然后,采用列联表分析、假设检验、回归分析等方法,对免费和付费在线问答社区用户参与行为特征和交互影响关系进行比较分析;最后,指出免费和付费在线问答社区存在的突出问题,并提出优化对策和建议。

2 研究综述

2.1 在线问答社区的相关研究

2.1.1 免费在线问答社区

免费在线问答社区普遍采取的是社会化问答模式,该模式下他们主要通过社区的方式构建信息交换渠道[24],因此又被称为社会化问答社区(Social Q&A Community),Yahoo! Answers、Quora、StackExchage 是该类型社区的典型代表。目前,研究者从用户需求、回答方式和内容质量等方面明确了免费在线社区用户参与行为的特征。在用户需求方面,E. Choi 等指出社会化问答社区适合满足用户对观点、想法和推荐信息的需求,而协同问答和专家问答适合满足用户对客观、事实信息的需求[25]。C. Shah 等指出用户使用社会化问答社区的目的是为了获得丰富的观察视角,参与更多的社交互动,以及获得更有针对性的内容。在信息分享方面,观点型、建议型和社交型问题通常可以获得更多的用户回复[26]。在内容质量方面,M. R. Morris 等认为社会化问答社区能够提供更具个性化、多样化和启发性的观点[27]。C. Shah 和 V. Kitzie 认为相比参考服务社会化问答社区在信息生产的成本、数量、速度、参与和协作等方面更具优势,但是在准确性、权威性等方面存在不足[28]。F. M. Harper 等发现社会化问答社区生成的内容具有较高的多样性和响应性,但是生成内容的质量参差不齐[29]。

2.1.2 付费在线问答社区

值乎、分答等是我国出现的新一代付费在线问答社区,与早期的付费问答社区 Goolge Answers 相比,他们虽然同样采取专家问答模式,但是前者融入了更多的社会化元素,在定价规则、交互形式和激励机制等方面也进行了重新的设计。在定价规则方面,Google Answers 等传统付费在线问答社区采取提问者定价的模式,由于缺少对专家背景、能力的认识和了解,用户很难设置合适的提问价格[30]。值乎等新一代付费在线问答社区采取回答者定价的模式,专家可以结合个人能力和盈利需求设定价格,在保障个人收益的同时也简化了社区的定价过程。在激励机制方面,Google Answers 采取专家、平台按比例分配收益的方法,更高的收益占比降低了专家的社区声誉[31]。知乎、值乎等社区允许用户支付较少的费用“偷听”回答内容,“偷听”所得收益由问答双方平分,这种方式同时为问答双方带来了金钱收益[6]。在交互对象和交互形式方面,Google Answers 由平台指定回答者,问答双方只能通过文字进行交流[18]。值乎等平台允许提问者自主选择回答者,问答双方可以通过文字、图片、语音等多种方式进行互动[23]。

2.2 用户参与行为的相关研究

2.2.1 用户参与行为的类型

在线问答社区提供了丰富的信息交互场景,允许用户通过提问、回答、评价等方式参与社区活动,从角色上可以将他们分为提问者、回答者和评价者。提问者通过提问发起与社区同行或付费专家的讨论,提问类型通常反映了他们的信息需求。E. Choi 等[25]和 S. Kim 等[32]将用户提问类型分为搜索信息、获得建议、寻找观点和非信息,M. Harper 等将提问类型分为搜索信息和寻求对话两种类型[33],M. J. Blooma 将在线问答社区信息分为元认知信息、过程性信息和概念信息等[34],基于上述分类研究对在线问答社区用户的搜索目标、回答质量和评价方式等进行了研究。回答者通过分享个人观点、看法、建议等参与社区活动,内、外部动因是激励他们分享的重要诱因。对于免费在线问答社区,利他、移情是激励回答者分享的重要内部动因,信誉、互惠等是关键的外部动因[17,35-38];对于付费在线问答社区,价格是影响专家分享的重要因素,更高的价格可以提高他们的回答长度、回答数量和存档价值[29,39],获取金钱收益成为激励专家分享的重要诱因。除了经济刺激,社交满足在激励专家与用户持续交互上同样发挥着重要的作用,其对金钱激励具有调节和催化的作用[40]。评价者根据个人经验对社区生成内容进行评价,不同的专业背景导致他们通常采取“个性化”的评价标准,需求类型、认知水平、情感体验和社交环境等均会影响他们对内容质量的判断[28,32,41]。

2.2.2 用户参与行为的交互影响

在线问答社区环境下，用户在特定信息需求指引下参与社区活动，通过信息交互实现对搜索内容的协同生产和协同过滤，面对差异化的信息需求、交互对象和社区环境他们会选择不同的交互策略，参与者之间会形成独特的交互影响关系[37]。在提问者与回答者的交互过程中，用户的提问方式会影响其他用户的回答策略，S. Kanthawala 等发现事实型、政策型和价值型问题在获得回答的数量、相关性和有效性方面存在着显著差异[41]。F. M. Harper 等发现不同主题用户会形成独特的交互偏好，主题和问题类型显著影响他们获得回答的数量和质量[29]。A. Y. K. Chua 和 S. Banerjee 发现问题类型和回答者能力显著影响回答质量和回答速度[42]。在提问者、回答者与评价者的交互过程中，C. Lampe 等发现不同类型发帖在获得回答数量、回答速度、评论和点赞数量等方面存在显著差异[43]。R. Gazan 基于累积优势理论指出回答－评价交互过程中存在着先动优势，先回答的用户通常可以获得更多的评价[44]。S. C. J. Sin 等发现更多的提问者参与和更为复杂的交互过程可以提高用户获得回答的质量，因此研究建议可以通过优化协商过程提高用户信息搜索的效果[45]。

综上所述，在线问答社区独特的信息交互模式为用户提供了多种类型的社区参与形式，通过发布需求、分享经验和参与讨论共同完成知识的分享、过滤和传播。研究者对在线问答社区用户信息需求、参与动因和交互影响关系的研究，一定程度上揭示了社区用户共同的行为规律，然而对于不同模式在线问答社区形成的独特用户行为方式尚缺少足够的研究。对此，本文主要从用户的信息需求、参与特征和交互关系三个方面对免费和付费在线问答社区用户的行为特征和交互策略进行比较研究，明确了用户访问免费和付费在线问答社区的目的，揭示两类社区用户独特的参与方式和差异化的交互影响关系。

3 研究方法

3.1 数据获取

3.1.1 数据来源

知乎是目前国内少数集免费和付费为一体的在线问答社区，本文选择其作为研究对象主要基于以下两个方面的考虑：一是自 2010 年创立以来知乎经历了快速的发展，积累了庞大的免费和付费用户群体，是当前主流的在线问答社区；二是知乎同时构建了免费、付费两种在线问答社区，他们分别采用了社会化问答和专家问答模式，选择对他们进行比较研究具有更好的代表性。

3.1.2 爬取目标

本研究需要分别爬取免费和付费在线问答社区数据，内容涉及用户提问、回答、评价等多项指标。主题是影响在线问答社区用户参与的重要因素，因此本研究分别从多个主题爬取问答数据。由于不同主题用户数量差异较大，统计分析时可能因方差不齐而影响结果的准确性，因此研究选择从电影、心理学、游戏和医学这 4 个用户较多的主题爬取数据，这些主题共包括 1 891位付费回答者，占回答者总数的52.4%，该样本量具有足够的代表性，能够满足研究需要。

3.1.3 爬取过程

本研究首先采用火狐浏览器对知乎网站进行抓包分析，以获取免费和付费问答数据的 API，然后通过 Python 编程爬取相关数据，最后采用 MongoDB 数据库对数据进行存储。详细的数据爬取过程如下：①针对付费问答数据，首先爬取 1 891 位付费回答者的个人 ID；然后从每个主题随机抽取 150 位回答者，并爬取他们的历史回答数、提问价格、问题内容、回答发布时间、回答时长和评价数等信息；最后，从爬取的问题中随机抽取 600 条（每个主题 150 条）进行人工编码。②针对免费问答数据，首先从话题广场搜索上述四个主题的问题，并爬取所有问题的内容、发布时间和 ID；然后，爬取每个问题的回答数量、回答发布时间、回答内容、评价数和关注数等信息；最后，从爬取的问题中选择 600 条（每个主题 150 条）进行人工编码。如图 1 所示：

爬虫设计： 抓包分析——>API 设计——>Python 编程爬取知乎 JSON 数据

↓

免费数据爬取： 从话题广场选定主题——>每个主题爬取 150 个问题——>爬取每个问题的前 20 条回答及评价信息——>从中随机选取 600 条问题准备编码；

付费数据爬取： 从付费主页选定主题——>从每个主题随机抽取 150 位回答者——>获取每位回答者前 20 条问题及回答、评价信息——>从中随机选取 600 条问题准备编码。

↓

数据整理： 为便于分析，需要对爬取的问题内容、回答内容等数据进行处理。如：按照编码类型对问题编码，通过 Python 生成免费提问和回答的长度，根据语音时长计算付费回答长度，通过 Python 对问答发布时间进行格式转换，对付费回答者和问题所属的主题进行审核等。

数据分析： 卡方检验、Kolmogorov-Smirnov 非参数检验、Kruskal-Wallis 检验、负二项回归分析。

图 1 数据获取流程

3.2 数据处理

由图 1 可见，爬取的免费和付费问答数据包括很多非结构化数据，由于无法直接进行分析，所以需要进

行预处理,详细的数据处理过程如下。

3.2.1　问题类型编码

本文采用内容分析法对免费和付费问题类型进行编码。首先,研究基于 S. Kim 等[29] 和 F. M. Harper 等[32] 的分类标准提出由事实型、建议型和观点型构成的预编码框架。然后,选择前 20% 的免费、付费问题进行预编码,通过预编码研究发现上述三种问题类型无法满足所有数据的分类需要。对此,研究引入 M. R. Morris 框架中的推荐型、自我表现型和社交型问题[15],并将推荐型问题细分为主观推荐和客观推荐两种类型,研究最终确定的编码框架包括 7 种问题类型,详细定义和样例如表 1 所示。最后,选择两位编码者对所有问题进行编码,通过交叉信度检验,编码结果的一致性达到了 84.5%,对于编码不一致的问题,经过讨论最终确定分类。

表 1　问题类型定义与样例

问题类型	定义与样例
事实型	定义:指针对特定事实、现象的标准回答或客观解释。 样例:《辐射》是 Fallout 的正确翻译么?
建议型	定义:指寻求他人的建议或可行的解决方案,重在解决问题,而非调查观点。 样例:如何保持在 1 - 2 小时内专注地做一件事情?
观点型	定义:指调查他人的想法、品味或引起讨论。 样例:请问现在德国医生的缺口大吗,未来两三年内拿到 offer 的可能性有多大呢?
客观推荐型	定义:指设定客观、具体的推荐标准,并要求他人按照标准进行推荐。 样例:有哪些心理学入门的书推荐?
主观推荐型	定义:指设定模糊、主观的推荐标准,允许他人根据主观偏好进行推荐。 样例:人们对生物有哪些常见的误解?
自我表现型	定义:指希望通过提问表达自己的思想、观点和看法,希望获得他人认同而非回答。 样例:一旦和身边的人处不好关系就会非常难受,很痛苦想改变!
社交型	定义:指希望与回答者建立特定的关系。 样例:我已经被你的人格魅力控制住了,做个自我介绍吧。

3.2.2　变量处理

本文爬取数据中包含一定的非结构化数据,需要将其转化为结构化数据才能进行分析。由于免费问题、免费回答和付费问题爬取的都是文本数据,所以研究采用 Python 编程统计了上述内容的字数,并将其作为免费问题长度、免费回答长度和付费问题长度保存到数据库。由于付费问答采取语音回答的方式,所以研究首先随机抽取 100 条付费语音回答,收听并统计上述回答的平均语速,然后将平均语速与回答时长相乘从而得到付费回答的长度。对于问题主题,免费问题都有统一的主题 ID,付费问题只有对应的答主 ID,所以研究主要依靠答主判断付费问题的主题,由于付费在线问答社区中存在着一位答主属于多个主题的情况,对此研究将根据问题内容确定其所属主题。在数据分析时,为了控制问题和回答发布时间对用户参与频次的影响,研究通过 Python 编程的方式统计从免费问题、免费回答和付费问题发布到数据爬取间隔的天数,作为免费问题发表时间、付费问题发表时间和免费回答发表时间变量,由于付费在线问答社区规定必须在两日内回复,所以不再统计付费回答时间。

3.3　数据分析

3.3.1　研究变量　本文共涉及 15 个研究变量,这些变量通过自动爬取、人工编码和程序统计等多种方式获得,详细的变量定义和处理方式如表 2 所示:

表 2　研究变量信息

序号	变量名称	数据内容	获取方式
1	免费问题类型	免费提问所属的信息需求类型	人工编码
2	免费问题主题	免费问题所属的主题类型	自动爬取
3	免费问题长度	免费问题包含的字符数	程序统计
4	免费问题发布时间	从问题发布到数据爬取中间间隔的天数	程序统计
5	免费回答长度	免费回答包含的字符数	程序统计
6	免费回答发布时间	从回答发布到数据爬取中间间隔的天数	程序统计
7	免费评价数	免费回答获得的点赞数量	自动爬取
8	免费关注数	免费问题获得的关注数量	自动爬取
9	付费问题主题	付费问题所属的主题类型	自动爬取
10	付费问题类型	付费提问所属的信息需求类型	人工编码
11	付费问题长度	付费问题包含的字符数	程序统计
12	付费问题发布时间	从问题发布到数据爬取中间间隔的天数	程序统计
13	历史回答数量	回答者总的付费回答数量	自动爬取
14	付费回答长度	付费回答包含的字符数	程序统计
15	付费评价数	付费回答获得觉得值的数量	自动爬取

3.3.2　数据分析　本文从信息需求、参与特征和交互影响关系三个方面对免费和付费在线问答社区用户的参与行为进行比较分析,首先通过描述统计分析问题类型、问答长度、回答数量和评价数量的分布规律;然后通过列联表分析、非参数检验对两类社区用户的提问数量、回答数量和评价数量等进行比较;最后通过负二项回归方法对两类社区用户在提问 - 回答、提问 - 回答 - 评价之间的交互影响关系进行比较分析。

4　研究结果

4.1　免费和付费问题类型的比较

4.1.1　免费和付费问题类型的分布

为明确免费和付费在线问答社区用户信息需求上的差异,研究采用卡方检验对两类社区问题类型的分

布情况进行比较,统计结果显示 Pearson 卡方值 = 134.279、P < 0.001,Tau 系数 = 0.029、P < 0.001,上述结果表明两类社区问题类型的分布存在显著差异。由图 2 可见,免费在线问答社区观点型问题的数量最多,占比达到了 51%,其次是主观推荐型问题,出现频次最少的是自我表现型问题,占比为 2.5%,由于知乎的免费在线问答社区采取推荐回答者的方式,用户事先并不知道具体的回答对象,所以没有出现社交型问题。付费在线问答社区建议型问题最多,占问题总数的 35.3%,其次是观点型问题,占问题总数的 30.3%,数量最少的是社交型问题,占比为 3.2%。

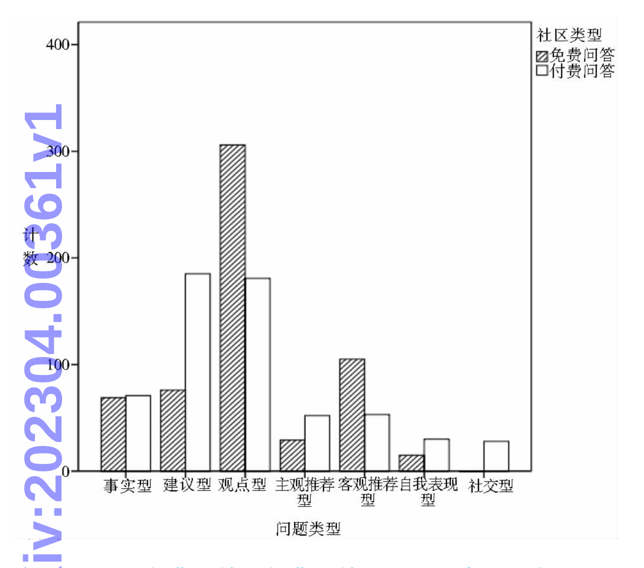


图 2　免费问答和付费问答社区问题类型分布

4.1.2　主题与免费、付费问题类型

主题是影响社区用户交互的重要情境因素[19],为比较不同主题下免费和付费问题类型的分布差异,研究采用多重列联表分析检验不同主题下社区类型与问题类型的关系。统计结果显示电影、心理学、游戏和医学 4 个主题的 Pearson 卡方值和显著性分别为 Pearson 卡方电影 = 67.009***、Pearson 卡方心理学 = 47.893***、Pearson 卡方游戏 = 39.876***、Pearson 卡方医学 = 12.253***,上述结果表明 4 种主题免费和付费问题类型的分布均存在显著差异。Tau 和列联系数分别为 Tau 电影 = 0.068、列联系数电影 = 0.427,Tau 心理学 = 0.04、列联系数心理学 = 0.371,Tau 游戏 = 0.034、列联系数游戏 = 0.343,Tau 医学 = 0.013、列联系数医学 = 0.198,这一结果表明电影主题免费和付费问题类型的分布差异最为明显,医学主题的分布差异最不明显,见图 3、图 4。

4.2　免费和付费参与行为的比较

4.2.1　提问、回答和评价特征的比较　研究采用 Kolmogorov-Smirnov 非参数检验对免费和付费在线问答社区的问题长度、回答长度和评价数量进行了比较。由表 3 可见,免费和付费在线问答社区中用户的问题长度、回答长度和评价数量均存在显著差异。其中,付费问题长度的均值为 48.81 个字符,显著高于免费问题;免费回答长度的均值为 802.61 个字符,显著高于付费回答;免费回答获得评价数的均值为 548.45 个,显著高于付费回答获得的评价数均值。

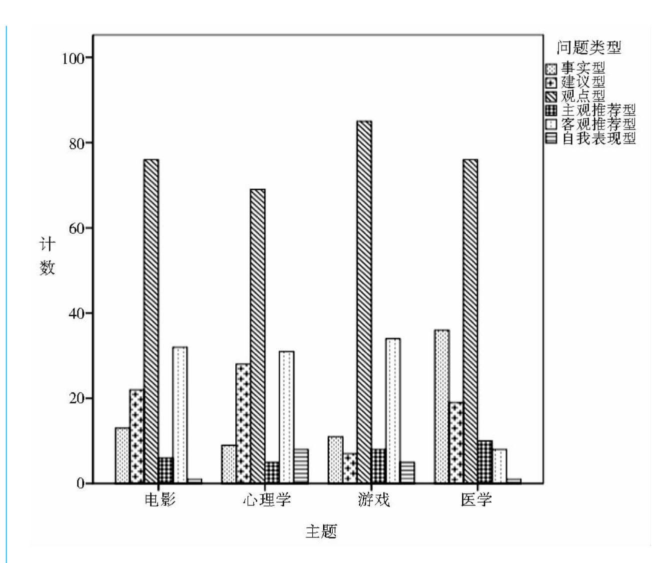


图 3　免费问答社区中分主题问题类型分布

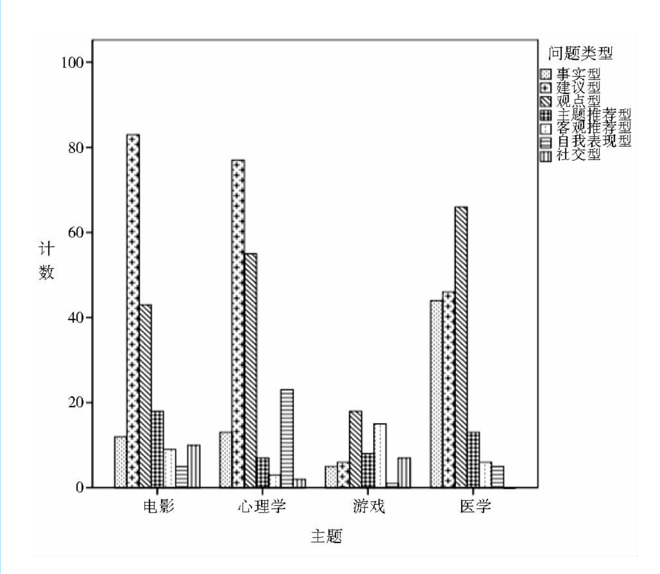


图 4　付费问答社区中分主题问题类型分布

表 3　在线问答社区类型与提问、回答和评价特征

特征	社区类型	均值（单位:字符）	方差	Z	P
问题长度	免费	22.37	8.562	12.03	0.00
	付费	48.81	14.613		
回答长度	免费	802.61	949.762	8.92	0.00
	付费	317.48	123.547		
评价数量	免费	548.45	1 466.114	10.75	0.00
	付费	1.89	4.288		

4.2.2 免费回答数和关注数的比较　由于每个付费问题只有一位回答者，所以本文主要对免费在线问答社区中用户的回答和关注数进行分析。Kruskal-Wallis 检验结果显示不同问题类型和主题下，免费回答数和关注数均存在显著差异（不同问题类型的卡方回答数量 = 34.93***、卡方关注数量 = 40.27***；不同主题的卡方回答数量 = 51.9***、卡方关注数量 = 21.37***）。由图 5a 可见，自我表现型和主观推荐型问题能够获得回答的数量较多，事实型问题获得的回答最少。由图 5b 可见，心理学主题获得的回答数量最多，最少的是医学主题。由图 5c 可见，主观推荐型问题受到用户关注的数量最高，而事实型问题获得关注的数量最少。如图 5d 可见，心理学主题关注的人数最多，而游戏主题关注的人数最少。

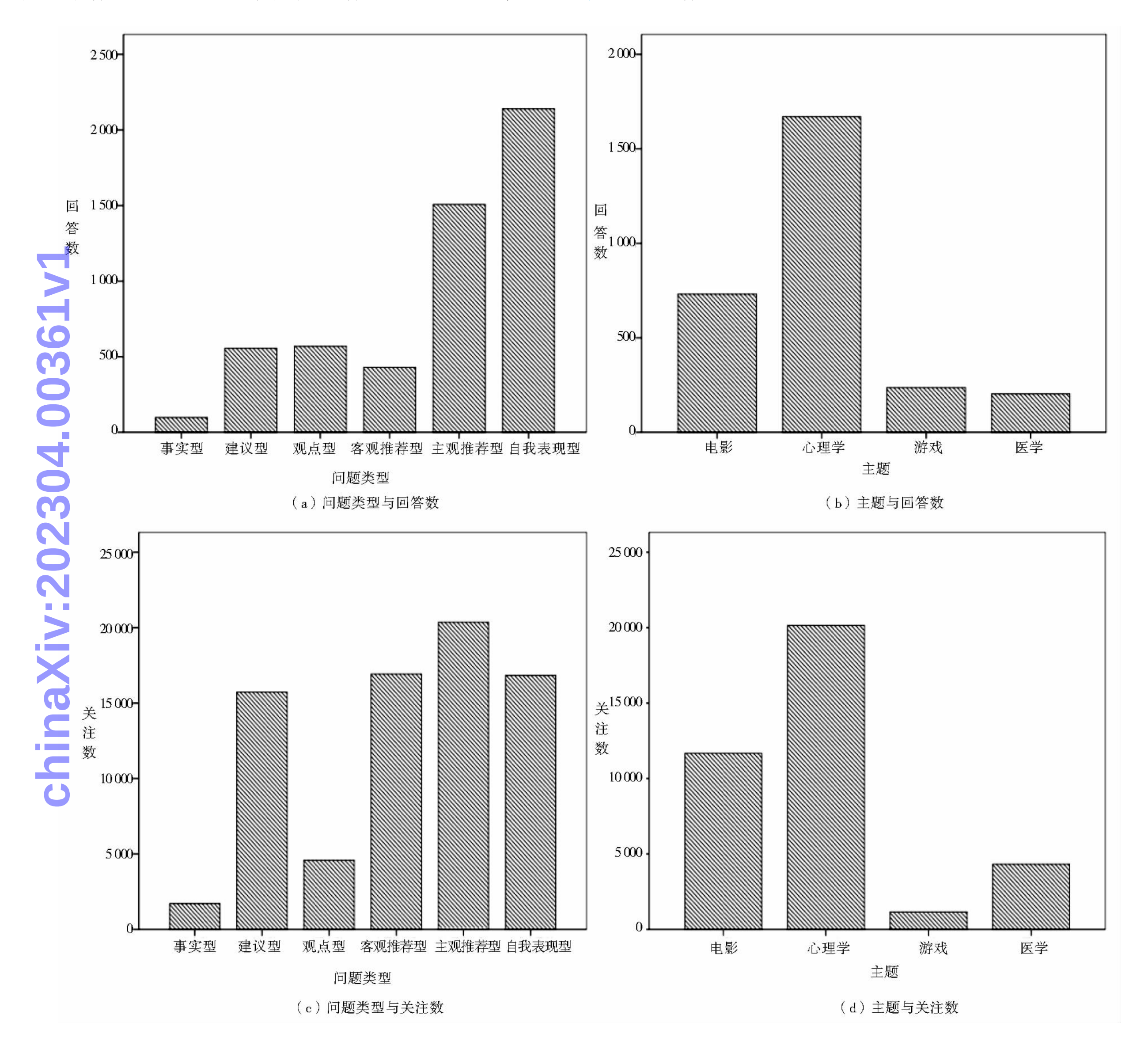


图 5　免费问答社区中问题类型、主题与回答数和关注数

4.3 用户交互影响关系的比较

4.3.1 研究模型的构建　为比较免费和付费在线问答社区中，提问者、回答者和评价者之间形成的独特交互规律，本文选择问题长度、问题类型、回答长度、主题作为自变量，问题发布时间和回答发布时间作为控制变量，分别分析上述变量对免费回答长度、回答数量和评价数量的影响。将历史回答数、价格、问题长度、问题类型、回答长度、主题作为自变量，问题发布时间作为控制变量，分别分析上述变量对付费回答长度和评价数量的影响。在回归模型构建过程中，研究对问题类型、主题等分类变量进行了虚拟化处理，为了避免“虚拟变量陷阱”导致的共线性问题，免费问题选择观点型问题为基准组，付费问题选择建议型问题为基准组，主题变量以“电影”主题为基准组，通过 Vuong 检验

确定采用稳健标准差负二项式回归分析，最终的检验模型如下：

免费问答：

$$Y_{\text{回答长度}}=\beta_0+\beta_1X_{\text{问题长度}}+\beta_2X_{\text{事实型}}+\beta_3X_{\text{建议型}}+\beta_4X_{\text{主观推荐型}}+\beta_5X_{\text{客观推荐型}}+\beta_6X_{\text{自我表现型}}+\beta_7X_{\text{心理学}}+\beta_8X_{\text{游戏}}+\beta_9X_{\text{医学}}+\beta_{10}X_{\text{问题发布时间}}\quad\text{（公式1）}$$

$$Y_{\text{回答数量}}=\beta_0+\beta_1X_{\text{问题长度}}+\beta_2X_{\text{事实型}}+\beta_3X_{\text{建议型}}+\beta_4X_{\text{主观推荐型}}+\beta_5X_{\text{客观推荐型}}+\beta_6X_{\text{自我表现型}}+\beta_7X_{\text{心理学}}+\beta_8X_{\text{游戏}}+\beta_9X_{\text{医学}}+\beta_{10}X_{\text{问题发布时间}}\quad\text{（公式2）}$$

$$Y_{\text{评价数量}}=\beta_0+\beta_1X_{\text{问题长度}}+\beta_2X_{\text{回答长度}}+\beta_3X_{\text{事实型}}+\beta_4X_{\text{建议型}}+\beta_5X_{\text{主观推荐型}}+\beta_6X_{\text{客观推荐型}}+\beta_7X_{\text{自我表现型}}+\beta_8X_{\text{心理学}}+\beta_9X_{\text{游戏}}+\beta_{10}X_{\text{医学}}+\beta_{11}X_{\text{问题发布时间}}+\beta_{12}X_{\text{回答发布时间}}\quad\text{（公式3）}$$

付费问答：

$$Y_{\text{回答长度}}=\beta_0+\beta_1X_{\text{历史回答数}}+\beta_2X_{\text{价格}}+\beta_3X_{\text{问题长度}}+\beta_4X_{\text{事实型}}+\beta_5X_{\text{建议型}}+\beta_6X_{\text{主观推荐型}}+\beta_7X_{\text{客观推荐型}}+\beta_8X_{\text{自我表现型}}+\beta_9X_{\text{心理学}}+\beta_{10}X_{\text{游戏}}+\beta_{11}X_{\text{医学}}+\beta_{12}X_{\text{问题发布时间}}\quad\text{（公式4）}$$

$$Y_{\text{评价数量}}=\beta_0+\beta_1X_{\text{历史回答数}}+\beta_2X_{\text{价格}}+\beta_3X_{\text{问题长度}}+\beta_4X_{\text{回答长度}}+\beta_5X_{\text{事实型}}+\beta_6X_{\text{建议型}}+\beta_7X_{\text{主观推荐型}}+\beta_8X_{\text{客观推荐型}}+\beta_9X_{\text{自我表现型}}+\beta_{10}X_{\text{心理学}}+\beta_{11}X_{\text{游戏}}+\beta_{12}X_{\text{医学}}+\beta_{13}X_{\text{问题发布时间}}\quad\text{（公式5）}$$

4.3.2　分析结果　免费回答长度的回归分析结果显示（见表4），免费问题长度显著负向影响回答长度，除事实型问题外其他类型问题的回答长度均显著高于观点型问题，心理学、游戏和医学主题下的回答长度显著低于电影主题。免费回答数量的回归分析结果显示，问题长度显著负向影响回答数量，事实型问题对回答数量的影响显著低于观点型问题，心理学主题的回答数量显著高于电影主题，游戏和医学则显著低于电影主题。免费评价数量的回归分析结果显示，问题长度显著负向影响评价数量，事实型问题的评价数量显著低于观点型问题，回答长度显著正向影响评价数量，心理学问题的评价数量显著高于电影主题，游戏和医学则显著低于电影主题。

表4　免费问答社区负二项回归分析结果

变量名	回答长度		回答数量		评价数量	
	Coef.	P	Coef.	P	Coef.	P
自变量						
问题长度	-0.014 1	0.000 0	-0.044 4	0.000 0	-0.042 7	0.000 0
回答长度	—	—	—	—	0.000 6	0.000 0
事实型	-0.433 7	0.000 0	-1.555 8	0.000 0	-1.502 0	0.000 0
建议型	0.151 9	0.017 0	-0.426 5	0.186 0	0.041 8	0.808 0
主观推荐型	0.392 2	0.000 0	0.120 8	0.749 0	0.038 5	0.807 0
客观推荐型	0.380 1	0.000 0	1.069 8	0.000 0	1.046 5	0.000 0
自我表现型	-0.281 3	0.007 0	1.074 7	0.078 0	1.743 0	0.000 0
心理学	-0.526 3	0.000 0	0.806 9	0.008 0	0.463 8	0.000 0
游戏	-0.566 1	0.000 0	-0.949 1	0.000 0	-1.265 1	0.000 0
医学	-0.720 7	0.000 0	-1.029 6	0.000 0	-0.640 0	0.000 0
控制变量						
问题发布时间	0.000 4	0.000 0	0.000 6	0.000 0	0.000 2	0.026 0
回答发布时间	—	—	—	—	0.000 6	0.000 0
cons	7.159 9	0.000 0	6.857 8	0.000 0	6.081 0	0.000 0
卡方值	638.040 0	0.000 0	334.830 0	0.000 0	1 279.740 0	
Log Likelihood	-54 396.977 0		-3 933.945 4		-40 287.117 0	

付费回答长度的回归分析结果显示（见表5），历史回答数对回答长度有显著的正向影响，有经验的回答者倾向于分享更长的内容；价格对回答长度的影响不显著；问题长度显著正向影响回答长度，详细的问题表述能够激励回答者的知识分享；除事实型问题的回答长度显著低于建议型问题外，其他问题类型的回答长度均与建议型问题相当；游戏、医学主题的回答长度显著低于电影主题，心理学与电影主题的回答长度不存在显著差异。付费评价数量的回归分析结果显示，历史回答数、价格和回答长度显著正向影响评价数量；除客观推荐型问题的评价数量显著低于建议型问题外，其他类型问题与建议型问题的评价数量相当；心理学、游戏和医学主题的评价数量均显著低于电影主题。

表5　付费问答社区负二项回归分析结果

变量名	回答长度		评价数量	
	Coef.	P	Coef.	P
自变量				
历史回答数	0.000 9	0.000 0	0.000 6	0.253 0
价格	0.000 8	0.336 0	0.008 4	0.002 0
问题长度	0.004 6	0.002 0	-0.010 2	0.028 0
回答长度	—	—	0.002 6	0.000 0
事实型	-0.150 7	0.016 0	-0.065 6	0.705 0
观点型	-0.022 1	0.563 0	0.025 1	0.860 0
主观推荐型	0.044 9	0.394 0	-0.230 2	0.155 0
客观推荐型	-0.010 3	0.920 0	1.115 4	0.000 0
自我表现型	-0.073 5	0.329 0	0.248 0	0.217 0
社交型	-0.021 2	0.842 0	0.329 4	0.268 0
心理学	-0.067 1	0.125 0	-0.519 0	0.001 0
游戏	-0.133 3	0.042 0	-0.691 3	0.001 0
医学	-0.094 5	0.039 0	-0.810 0	0.000 0
控制变量				
问题发布时间	—	—	0.003 2	0.000 0
cons	5.536 1	0.000 0	-0.256 7	0.432 0
卡方值	101.180 0	0.000 0	197.800 0	0.000 0
Log Likelihood	-3 820.527 9	-977.397 4		

5　讨论与启示

通过上述分析，本研究发现免费和付费在线问答用户在信息需求、参与方式和交互影响关系上存在着诸多差异，下面主要结合统计分析结果进行讨论，并提出具体的系统优化建议。

5.1　用户信息需求和提问方式的比较

首先，免费和付费在线问答社区用户的信息需求存在显著差异，对于免费问答用户而言，社会化的问答模式为他们提供了丰富的调查对象[46]，能够帮助他们获得多样的观点和看法[16]，因此他们提出了更多的观点型问题；对于付费问答用户而言，他们提问的目的是

希望获得专家的建议和指导，他们追求回答的客观性而非多样性[47]，因此他们提出了更多的建议型问题。其次，对问题长度的分析结果显示，免费和付费问答用户会选择不同的提问策略[37]，免费提问者为了吸引更多用户回答会选择抽象、概括的问题[25]，付费提问者为了获得客观回答通常会提供更为详细的背景信息，这导致他们的提问内容更长。

基于上述分析，本研究认为免费在线问答社区可以开设专门的社会调查栏目，增加系统推荐答主的上限，从而更好地满足用户对多样化观点的调查需求。另外，针对事实型问题回答和关注数量偏少的问题，社区应当加强对这类问题的奖励和推广力度，同时提高答主推荐的准确性。对于付费在线问答社区，研究认为应当进一步优化社区的问题发布功能，使用户能够快捷的完成图片、视频和链接信息的发布。另外，社区可以提供智能化的辅助提问功能，帮助用户更清楚、准确地描述个人需求。

5.2 用户回答方式的比较

在回答数量上，免费问题通常可以获得多个用户的回答，回答内容具有较高的多样性，付费问题由于是面向特定专家提问，所以只能获得一种回答。其次，在回答方式上，免费回答的长度显著高于付费回答，这是因为免费在线问答社区没有回答字数和形式的限制，他们可以通过图片、链接等多种方式表达个人观点[23]。付费在线问答社区严格限定了回答的方式和时间，语音问答方式虽然提高了用户的社会临场感，但同时也对专家凝练观点的能力提出了更高要求[22]。在交互方式上，免费在线问答社区中，问题长度与回答长度存在显著地负向影响关系，由于没有情境的限制回答者可以更为自由的表达观点，因此简洁、概括和开放式的问题可以吸引更多用户的回答。付费在线问答社区中，问题长度与回答长度存在显著的正向影响关系，详细、具体的问题描述能够帮助专家更准确的做出判断，因此回答的内容更长[29]。

基于上述分析，研究认为免费回答的内容通常更长、更复杂，因而需要花费更多的时间和精力，对此免费在线问答社区应当提供更方便、快捷的内容编辑功能，提高用户内容创作的效率。另外，针对免费回答者对开放式问题的偏好，社区应当提供更多措施激励用户分享其他类型的问题。对于付费在线问答社区，本研究认为应当取消对回答字数和次数的限制，允许问答双方自由选择文字、语音和视频等多种方式进行交互，通过增加交互频次、丰富交互方式提高付费问答质量。

5.3 用户评价方式的比较

在评价数量上，研究发现免费回答者获得的评价数显著高于付费回答者，一方面这是因为知乎的免费在线问答社区拥有更为庞大的用户基础，多种内容推荐方式极大地提高了问题的曝光度；另一方面，付费会降低用户对回答者的感激，同时提高他们对内容质量的预期，这也会导致用户对付费回答评价数量的降低[18]。另外，研究发现免费和付费两类社区的问题长度、回答长度分别负向和正向影响评价数，这表明无论是免费或是付费在线问答社区，用户都喜欢对简单、抽象的问题进行评价，此外更长的回答内容也可以吸引更多的用户评价。上述结果可能是因为简单、抽象的问题通常不涉及专业的知识背景，因而更容易被评价。另外，对于缺少专业知识的评价者，回答长度可能是他们评价信息质量的重要指标，更长的内容意味着更高的质量。

基于上述分析，本研究认为对于免费在线问答社区，系统应当提高内容推荐的准确性，确保被推荐的用户有兴趣、有能力对推荐内容进行评价。对于付费在线问答社区存在的用户评价数量低的情况，本研究认为应当丰富评价方式，除星级外还可以增加文字评价功能，从而方便用户完整地表达个人观点。

5.4 用户激励措施的比较

在激励措施上，免费回答可以帮助用户获得大量的社区成员关注，作为重要的结构资本提高他们在社区中的影响力[5]。之前有研究发现价格可以提高付费回答的数量和长度[48]，然而本研究结果显示价格对回答长度的影响并不显著，这可能是因为值乎等新兴付费问答社区普遍采取回答者定价模式，价格从一种激励措施转变为招揽用户的手段。同时，本研究也发现历史回答数和问题长度可以显著提高回答者分享的长度，这表明丰富的付费回答经验和明确的问题表述能够激励回答者的信息分享[32]。

基于上述分析，本研究认为无论是免费或是付费在线问答社区，都应当采取多种激励手段，从内部和外部两个方面引导用户进行分享。对于付费在线问答社区，应当借鉴免费在线问答社区的成功经验，设计针对付费问答用户的荣誉、勋章等虚拟奖励机制，以提高用户付费提问和付费分享的内在动机。对于免费在线问答社区，平台应当帮助免费问答用户变现之前积累的社会资本，推动免费收益向付费收益的转化。

6 结论

本研究采用内容分析、非参数检验和回归分析等方法分析了免费和付费在线问答社区用户的信息需求和参与特征，发现上述两类社区能够满足不同类型用户的信息需求，他们在用户参与特征和交互策略等方面存在着显著差异。从理论意义方面来看，本文丰富了在线问答社区用户行为理论研究，为探索用户参与在线问答社区的影响因素提供必要的支持。从实践意义方面来看，本文针对免费和付费两类社区提出的优化方案能够进一步突出两类社区的特点和优势，从而帮助他们实现差异化的发展策略。

本研究尚存在一定的局限性，首先，本文选择知乎作为研究对象，数据来源较为单一，其次，本文采用行为数据进行分析，缺少对用户主观意愿的调查和研究。针对该问题，后续笔者将采用深度访谈与问卷调查等方式，对免费和付费在线问答社区用户的搜索策略、分享意愿等进行调查，深入研究导致用户参与行为差异的主观原因。最后，针对本文提到的免费用户向付费用户的转化问题，笔者还将对付费提问者的决策过程和回答者的免费、付费分享意愿进行深入研究，结合免费和付费在线问答社区特点，实现两类社区的差异化和融合化发展。

参考文献：

[1] FENG Y Y, YE H. Why do you return the favor in online knowledge communities? a study of the motivations of reciprocity[J]. Computers in human behavior, 2016, 63: 342 – 349.

[2] YANG X Y, SONG S Q, ZHAO X S, et al. Understanding user behavioral patterns in open knowledge communities[J]. Interactive learning environments, 2018, 26(2): 245 – 255.

[3] JIN J H, LI Y J, ZHONG X J, et al. Why users contribute knowledge to online communities? An empirical study of an online social Q&A community[J]. Information & management, 2015, 52(7): 840 – 849.

[4] JIN X L M, ZHOU Z Y, LEE M K O, et al. Why users keep answering questions in online question answering communities: a theoretical and empirical investigation[J]. International journal of information management, 2013, 33(1): 93 – 104.

[5] 赵杨,袁析妮,李露琪,等. 基于社会资本理论的问答平台用户知识付费行为影响因素研究[J]. 图书情报知识,2018,184(4):17 – 25.

[6] 刘周颖,赵宇翔. 基于语音互动的付费知识问答社区运营模式初探——以分答和值乎为例[J]. 图书与情报,2017(4):38 – 46.

[7] GAZAN R. Microcollaborations in a social Q&A community[J]. Information processing & management, 2010, 46(6): 693 – 702.

[8] CHOUDHURY S, HARITH A. Exploring user behavior and needs in Q&A communities[C]// Proceedings of the European conference on social media ECSM 2014. Brighton: Academic Conferences and Publishing International, 2014: 80 – 89.

[9] EFRON M, WINGET M. Questions are content: a taxonomy of questions in a microblogging environment[J]. Proceedings of the American Society for Information Science & Tchnology, 2011, 47(1): 1 – 10.

[10] HARPER F M, MOY D, KONSTAN J A. Facts or friends? distinguishing informational and conversational in SQA[C]//International conference on human factors in computing systems. Boston: ACM, 2009: 759 – 768.

[11] DENG S, FANG Y, LIU Y, et al. Understanding the factors influencing user experience of social question and answer services[J]. Information research: an international electronic journal, 2015, 20(4): 1 – 17.

[12] FICHMAN P. A comparative assessment of answer quality on four question answering sites[J]. Journal of information science, 2011, 37(5): 476 – 486.

[13] CHANG Y P, DONG X B. Research on the impact of consumer interaction behaviour on purchase intention in an SNS environment: evidence from China[J]. Information development, 2014, 32(3): 1 – 13.

[14] JIN J H, YAN X B, WANG L, et al. Why users continually seek knowledge in online communities ? an empirical study of an online Q&A community[C]//2015 International conference on management science & engineering(22th). Dubai: IEEE, 2015: 19 – 22.

[15] MORRIS M R. What do people ask their social networks, and why? a survey study of status message Q&A behavior[C]//Proceedings of the 28th international conference on human factors in computing systems. Atlanta: ACM, 2010: 1739 – 1748.

[16] CHOI E, SHAH C. User motivations for asking questions in online Q&A services[J]. Journal of the Association for Information Science and Technology, 2015, 67(5): 1182 – 1197.

[17] JIN X L, LEE M K O, CHEUNG C M K. Understanding users' continuance intention to answer questions in online question answering communities[C]//The 9th international conference on e-lectronic business. Macau: USACO, 2009: 679 – 688.

[18] HSIEH G, KRAUT R E, HUDSON S E. Why pay?: exploring how financial incentives are used for question & answer[C]//Proceedings of the 28th international conference on human factors in computing systems. Atlanta: ACM, 2010: 1 – 10.

[19] WANG H X, Xi Z H, ZHANG J X, et al. An empirical study of financial incentivized question answering in social websites[C]//Proceedings of ACM workshop on hot topics in web systems and technologies. San Jose: ACM, 2017: 1 – 8.

[20] JAN S T K, WANG C, ZHANG Q, et al. Pay-per-question: towards targeted Q&A with payments[C]//Proceedings of the 2018

ACM conference on supporting groupwork. Sanibel Island: ACM, 2018: 1-11.

[21] ZHAO Y, LIU Z, SONG S. Why should I pay for the knowledge in social Q&A platforms? [C]// iConference 2018: Transforming digital worlds. Sheffield: Springer, 2018: 557-582.

[22] 赵宇翔,刘周颖,宋士杰. 新一代知识问答平台中提问者付费意愿的影响因素探究[J]. 数据分析与知识发现,2018,2(8):16-30.

[23] 张颖,朱庆华. 付费知识问答社区中提问者的答主选择行为研究[J]. 情报理论与实践,2018,41(12):25-30.

[24] KHUSRO S, ALAM A, KHALID S. Social question and answer sites: the story so far[J]. Program electronic library & information systems, 2017, 51(1): 170-192.

[25] CHOI E, KITZIE V, SHAH C. Developing a typology of online Q&A models and recommending the right model for each question type[J]. American Society of Information Science & Technology, 2012, 49(1): 1-4.

[26] SHAH C, KITZIE V, CHOI E. Modalities, motivations, and materials-investigating traditional and social online Q&A services[J]. Journal of information science, 2014, 40(5): 669-687.

[27] MORRIS M R, TEEVAN J, PANOVICH K. A comparison of information seeking using search engines and social networks[C]//Proceedings of the fourth international conference on weblogs and social media. Washington: AAAI, 2010: 613-617.

[28] SHAH C, KITZIE V. Social Q&A and virtual reference-comparing apples and oranges with the help of experts and users[J]. Journal of the Association for Information Science and Technology, 2012, 63(10): 2020-2036.

[29] HARPER F M, RABAN D, RAFAELI S, et al. Predictors of answer quality in online Q&A sites[C]//Conference on human factors in computing systems 2008. Florence: ACM, 2008: 865-874.

[30] EDLMAN B. Earnings and ratings at google answers [J]. Economic inquiry, 2012, 50(2): 309-320.

[31] JAFARI M, HESAMAMIRI R, SADJADI J, et al. Assessing the dynamic behavior of online Q&A knowledge markets: a system dynamics approach[J]. Program electronic library & information systems, 2012, 46(3): 341-360.

[32] KIM S, OH J S, OH S. Best-answer selection criteria in a social Q&A site from the user-oriented relevance perspective[J]. Proceedings of the American Society for Information Science & Technology, 2007, 44(1): 1-15.

[33] HARPER M, MOY D, KONSTAN J A. Facts or friends? distinguishing informational and conversational questions in social Q and A sites[C]//Proceedings of the SIGCHI conference on human factors in computing systems. Boston: ACM, 2009: 1-10.

[34] BLOOMA M J, KURIAN J C, CHUA A Y K, et al. Social question answering: analyzing knowledge, cognitive processes and social dimensions of micro-collaborations[J]. Computers & Education, 2013, 69(4): 109-120.

[35] KHANSA L, Ma X, LIGINLAL D, et al. Understanding members' active participation in online question-and-answer communities: a theory and empirical analysis[J]. Journal of management information systems, 2015, 32(2): 162-203.

[36] ZHAO L, DETLOR B, CONNELLY C E. Sharing knowledge in social Q&A sites: the unintended consequences of extrinsic motivation[J]. Journal of management information systems, 2016, 33(1): 70-100.

[37] ZHANG X, LIU S, CHEN X, et al. Social capital, motivations, and knowledge sharing intention in health Q&A communities[J]. Management decision, 2017, 55(7): 1536-1557.

[38] LOU J, FANG Y, LIM K H, et al. Contributing high quantity and quality knowledge to online Q&A communities[J]. Journal of the American Society for Information Science and Technology, 2013, 64(2): 356-371.

[39] BODOF D, RABAN D. Question types and intermediary elicitations[J]. Journal of the Association for Information Science & Technology, 2016, 67(2): 289-304.

[40] RABAN D R. The incentive structure in an online information market[J]. Journal of the Aamerican Society for Information Science & Technology, 2014, 59(14): 2284-2295.

[41] KANTHAWALA S, VERMEESCH A, GIVEN B, et al. Answers to health questions: Internet search results versus online health community responses[J]. Journal of medical Internet research, 2016, 18(4): 1-13.

[42] CHUA A Y K, BANERJEE S. So fast so good: an analysis of answer quality and answer speed in community question-answering sites[J]. Journal of the American Society for Information Science & Technology, 2013, 64(10): 2058-2068.

[43] LAMPE C, GRAY R, FIORE A T, et al. Help is on the way: patterns of responses to resource requests on facebook[C]//The 17th ACM conference on computer supported cooperative work & social computing. Baltimore : ACM, 2014: 3-15.

[44] GAZAN R. First-mover advantage in a social Q&A community [C]//The 48th Hawaii international conference on system sciences. Kauai: IEEE, 2015: 1616-1623.

[45] SIN S C J, LEE C S, THENG Y L. Social Q&A question-and-comments interactions and outcomes: a social sequence analysis[C]// ICADL 2016. Tsukuba: LNCS, 2016: 325-338.

[46] HARPER F M, WEINBERG J, LOGIE J, et al. Question types in social Q&A sites[J]. First Monday, 2010, 15(7): 1-20.

[47] HSIEH G, COUNTS S. Mimir: a market-based real-time question and answer service [C]// Proceedings of the 27th international conference on human factors in computing systems, CHI 2009. Boston: ACM, 2009: 1-10.

[48] JEON Y J, KIM Y M, CHEN Y. Re-examining price as a predictor of answer quality in an online Q&A site[C]//Sigchi conference on human factors in computing systems. Atlanta: ACM, 2010: 325-

328.

作者贡献说明：

齐云飞：收集研究数据，分析研究结果，撰写论文；
赵宇翔：提出研究思路，修正研究框架，指导数据分析；
刘周颖：辅助文献收集，审核校验论文；
孙晓宁：调整论文框架，指导论文撰写；
朱庆华：确定研究选题，指导研究设计，审核校验论文和定稿。

A Comparative Study on Users Information Interaction Behaviors Between Free and Fee Q&A Communities

Qi Yunfei[1] Zhao Yuxiang[2] Liu Zhouying[3] Sun Xiaoning[4] Zhu Qinghua[3]

[1] School of Information Management, Zhengzhou University of Aeronautics, Zhengzhou 450046

[2] School of Economics & Management, Nanjing University of Science & Technology, Nanjing 210094

[3] School of Information Management, Nanjing University, Nanjing 210023

[4] School of Information, Shanxi University of Finance & Economics, Taiyuan 030006

Abstract: [**Purpose/significance**] As the popular online Q&A mode, the free and fee online Q&A communities provide users with different information interaction scenarios, differentiated user needs can be met. This paper compares the participation behavior of two types of community users, and reveals the behavioral differences between them, finally some countermeasures and suggestions for the optimization of the two kinds of community service modes are put forward. [**Method/process**] The contingency table analysis, non-parametric test and negative binomial regression were used to analyze the interaction characteristics and interaction relations. [**Result/conclusion**] This paper finds that there are many differences between free and fee online Q&A users in information demand, participation characteristics and interaction relations, free online Q&A has the most opinion questions and paid online Q&A has the most advice questions; free questioners tend to ask abstract and general questions, while fee questioners tend to provide detailed background information; the length of questions negatively affects the number and length of free answers, and positively affects the length of fee answers; the question length and answer length positively affect the number of free and fee evaluations. This paper enriches the research on user participation behavior of online Q&A community, and provides some management suggestions for free and paid online Q&A community.

Keywords: online Q&A community payment-based Q&A platform pay for knowledge human information interaction